EDGEWOOD

CHEMICAL BIOLOGICAL CENTER

U.S. ARMY RESEARCH, DEVELOPMENT AND ENGINEERING COMMAND

ECBC-TR-616

AUTOMATED SPEECH INTELLIGIBILITY SYSTEM FOR HEAD-BORNE PERSONAL PROTECTIVE EQUIPMENT: PROOF OF CONCEPT

Karen M. Coyne
Daniel J. Barker

RESEARCH AND TECHNOLOGY DIRECTORATE

Jonathan P. Eshbaugh

Battelle
The Business of Innovation

BATTELLE MEMORIAL INSTITUTE
Columbus, OH 43201-2693

April 2008

ABERDEEN PROVING GROUND, MD 21010-5424

Disclaimer

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorizing documents.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
XX-04-2008	Final	Jan 2007 - Sep 2007

4. TITLE AND SUBTITLE
Automated Speech Intelligibility System for Head-Borne Personal Protective Equipment: Proof of Concept

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Coyne, Karen M.; Barker, Daniel J. (ECBC); and Eshbaugh, Jonathan P. (Battelle Memorial Institute)

5d. PROJECT NUMBER
7NBN2F

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
DIR, ECBC, ATTN: AMSRD-ECB-RT-PR, APG, MD 21010-5424
Battelle Memorial Institute, 505 King Avenue, Columbus, OH 43201-2693

8. PERFORMING ORGANIZATION REPORT NUMBER
ECBC-TR-616

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Law Enforcement Standards/NIST, Gaithersburg, MD 20899-8102
U.S. Department of Homeland Security, Science and Technology Directorate, Washington, DC 20528

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT-LIMIT 200 WORDS
An automated objective test system was developed to assess the impact of head-borne personal protective equipment on speech intelligibility and transmission. The system comprised talker and listener headforms, speech recordings, and speech recognition software. A recording of sentences from the Speech Perception in Noise test was transmitted from the speaker in the talker headform to microphones in the ears of the listener headform. The speech recognition software recorded the speech received at the listener headform. The responses were scored by hand. Speech transmission was assessed by separately placing two air-purifying respirators on the talker headform while speech intelligibility was evaluated by separately placing a chemical protective suit, an escape respirator with hood, and a ballistic helmet on the listener headform. A control condition with bareheaded talker and listener headforms was run for comparison. Speech transmission was statistically significantly degraded by both respirator conditions compared to the control condition. The protective suit and escape respirator both caused statistically significant decreases in speech intelligibility compared to the control and helmet conditions. These results show that an automated test system has been developed that can be used to quantify degradations in speech transmission and intelligibility when head-borne personal protective equipment is worn.

15. SUBJECT TERMS
Personal protective equipment; Speech transmission; Respirator; Speech intelligibility; Communications

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Sandra J. Johnson
U	U	U	UL	20	19b. TELEPHONE NUMBER *(include area code)* (410) 436-2914

Blank

PREFACE

The work described in this report was authorized under Project No. 7NBN2F. This work was started in January 2007 and completed in September 2007.

The use of either trade or manufacturers' names in this report does not constitute an official endorsement of any commercial products. This report may not be cited for purposes of advertisement.

This report has been approved for public release. Registered users should request additional copies from the Defense Technical Information Center; unregistered users should direct such requests to the National Technical Information Service.

Acknowledgments

The authors would like to thank the Office of Law Enforcement Standards, National Institute of Standards and Technology, and the Department of Homeland Security, Science and Technology Directorate for funding this project.

Blank

CONTENTS

FIGURES

TABLES

AUTOMATED SPEECH INTELLIGIBILITY SYSTEM FOR HEAD-BORNE PERSONAL PROTECTIVE EQUIPMENT: PROOF OF CONCEPT

INTRODUCTION

It is critical that first responders be able to communicate clearly with one another when responding to a chemical biological radiological and nuclear (CBRN) event. Head-borne personal protective equipment (PPE), such as respirators, hoods, and helmets, impacts speech intelligibility by interfering with speech transmission and reception. Current PPE standards address speech intelligibility while wearing a respirator but do not consider the impact of different chemical protective hood materials and thicknesses, or different helmet styles on speech reception.

The National Institute for Occupational Safety and Health (NIOSH) air-purifying respirator (APR) CBRN communications standard[3] assesses speech intelligibility during mask wear by scores resulting from the Modified Rhyme Test (MRT).[1] This is a subjective test that evaluates a listener's ability to identify single words spoken by a mask wearer. A major drawback of the MRT procedure is the time and cost required to conduct the test. In addition, the test participants must be attentive and motivated. Five speakers and three listeners are required to complete the test. Test speakers must speak without an accent and are required to maintain a sound output level between 75 and 85 dB when not wearing a mask and then duplicate the same vocal effort with a mask on. As the mask alters the sound level, there is no way to assess whether the words are being spoken with the same intensity. If all the speakers consistently speak at 75 dB, the performance rating of the mask might be lower than if an 85 dB sound level were maintained. Speakers may also over-enunciate words while wearing the mask, further altering the sound signal. While scores are averaged across speakers, it is possible that one very bad speaker could cause a respirator that should have passed to result in a failed test. Because this is a subjective test, results depend on the subject responses and cannot be reproduced exactly.

An automated objective test that predicts speech intelligibility based on sound quality parameters would be less expensive, faster, and independent of human subject performance. It would allow different mask/hood/voice projection unit combinations to be tested quickly and efficiently. The test system would require a talker headform with a speaker in the mouth cavity, a listener headform with microphones in the ears, recorded speech, and a mechanism for evaluating the speech signal received by the listener headform.

One option for evaluating the speech signal is a commercial off-the-shelf speech recognition software package. The speech intelligibility test used with this software may impact the results. Most speech recognition software packages are designed to use contextual clues to identify words. The MRT eliminates the use of contextual clues by using a carrier sentence with individual words. For example, "The word is sit." Additionally, it would be difficult to score the MRT test using the speech recognition software because the MRT uses a closed response set. That is, the listener has before him a set of 6 possible words from which to choose. For the above example, these choices would be: sit, sip, sill, sick, sin, and sing. The software has no such limitation. The MRT tests the ability to transmit leading or trailing consonants but does not test the vowel between the two. So, the software may identify the spoken word "sit" as "set". The two consonants were identified correctly, but the vowel was not. A human listener would have the closed response set before him and would likely identify the spoken word correctly if the two consonants were transmitted intelligibly.

Another speech intelligibility test is the Speech Perception in Noise (SPIN), which uses eight lists of 50 sentences.[2] For half the sentences, the response is highly predictable due to sentence context while responses for the other half have low predictability. An example of a high predictability sentence is "We saw a flock of wild geese." while a low predictability sentence is "Miss Black knew about the doll." The sentence is presented to the listener by a trained speaker and the listener writes the word that best completes the sentence. Scoring is simple: the answer is either correct or incorrect, accounting for spelling errors and homonyms.

The goal of the current effort was to develop an automated objective test system for quantifying the effects of respirators, hoods, and helmets on speech intelligibility and reception. This is the first step in developing a test method that could be utilized in a communications standard for first responder head-borne PPE.

2. METHODS

2.1 Test System

A test system was developed that included speech recognition software, recorded speech, a talker headform, and listener headform.

2.1.1 Speech Recognition Software

The Dragon Naturally Speaking speech recognition software package (Nuance Communications, Inc., Burlington, MA) was purchased and installed on a Pentium laptop computer. The laptop contained a SigmaTel C-Major Audio sound card (Austin, TX) and Sound Forge 8.0 (Sony Creative Software, Inc., Madison, WI) audio processing software. Two male volunteers "trained" the speech recognition software for their voices. The training involved reading selected passages from literature provided by the software to "familiarize" the speech recognition engine with the speaker's pronunciations, cadence, and inflections.

Trials were conducted using both the MRT and SPIN to determine which test would be better for assessing speech intelligibility using the software. Each of the two male speakers read two sets of 10 MRT words (List 1) and two sets of 10 high predictability SPIN sentences (Form 2.1) while bareheaded and while wearing an air-purifying respirator. The words and sentences used are provided in Appendix A. The sound level was measured using a sound level meter (Model 322, Center Technology Corporation, Taipei, Taiwan). The A-weighted fast response setting was used. For the MRT trials, the speaker was instructed to read the sentences at the same sound level throughout the sentence. For the SPIN trials, the speaker read the sentences in a normal speaking voice which naturally resulted in the last word of the sentence being softer than the first. Sound levels were recorded only for the last word. The speaker wore the hands-free AntiNoise ® PC Headset NC-91 (Andrea Electronics Corporation, Bohemia, NY) that came with the Dragon software. The background noise was approximately 46 dBA.

The MRT words were scored correctly if the computer recognized both the first and last consonant. That is, no penalty was assessed if the software identified "sit" as "set". As this modified version of the MRT did not include a closed response set, scores were not adjusted for guessing as they are for the traditionally administered MRT. The SPIN sentences were scored correct if the last word was identified correctly. No penalty was assessed if other words in the sentence were not correct.

2.1.2 Speech Recordings

One male native English speaking volunteer without a regional accent recorded the first SPIN list (Form 2.1) while wearing the NC-91 headset.[2] The volunteer was instructed to use normal inflections when speaking.

2.1.3 Talker Headform

A static rubber headform with a speaker inserted in the mouth cavity was used. An audio cable was connected from a desk top computer sound output (Intel 8280BA/BM AC97 Audio Controller, Santa Clara, CA) to a Kenwood U.S.A. stereo receiver (Model KR-AR080, Long Beach, CA). Speaker wire was then run to the sound inputs on the headform.

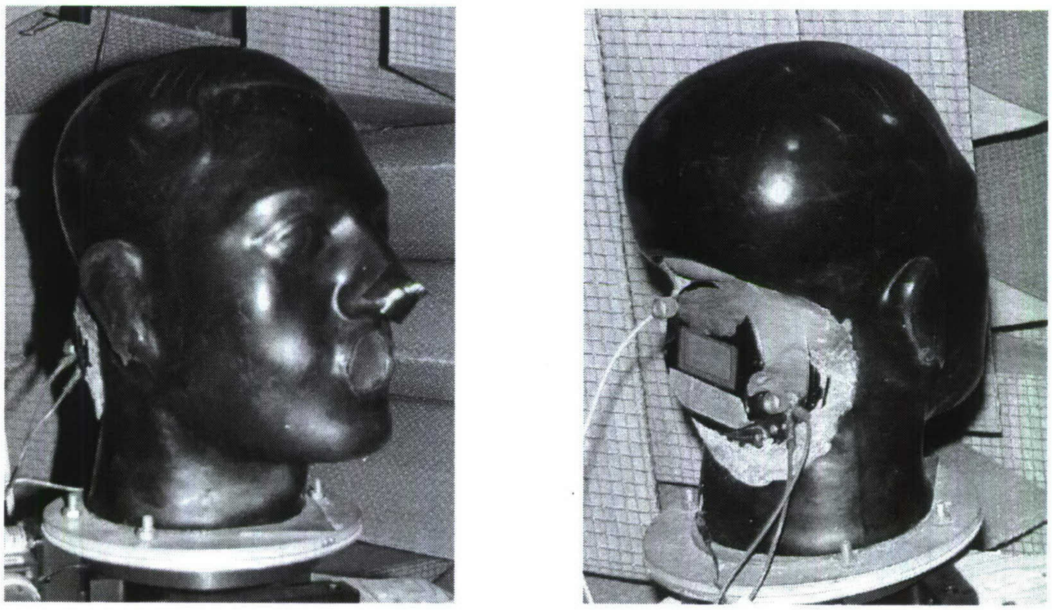

Figure 1. Talker Headform with Speaker Mounted in Back

2.1.4 Listener Headform

The Bruel and Kjaer (Naerum, Denmark) Type 4128 C head and torso simulator (HATS) was selected as the listener headform for this effort. The HATS has binaural sound quality microphones inserted into the ear canals and rubber pinnae that simulate ears.

The output from each of the HATS microphones was connected to a Bruel and Kjaer Sound Quality Conditioning Amplifier Type 2672. The left and right signal outputs were then connected by a custom-made cable to the sound card in the speech recognition computer. Only sound from the right microphone was transmitted to the sound card due to a limitation with the sound card.

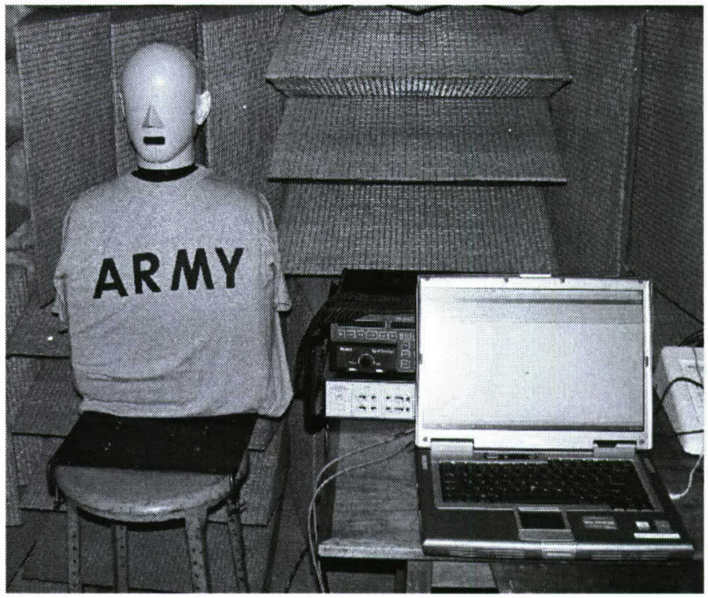

Figure 2. Listener Set Up Including Head and Torso Simulator, Amplifier, and Laptop

2.2 System Evaluation

The system was set up in an anechoic chamber (Eckel Industries, Inc., Cambridge, MA) with a headform separation distance of 2 m, ambient background noise (approximately 42 dBA), amplifier gain of 10; and recorded speech volume between 75 and 85 dBA. The recorded SPIN list was played through the talker headform. The HATS microphones received the audio signal and the right ear signal was sent to the sound conditioning amplifier and then to the laptop sound card. An MS Word document was used to display the speech identified by the speech recognition software. Five PPE and one control configuration were tested. Two of the conditions were chosen to test the impact of PPE on speech transmission while the other three configurations tested the impact on speech intelligibility. Two NIOSH-certified APRs, one protective hood, one ballistic protective helmet, and one escape respirator were used. The APRs used were the Mine Safety Appliance (MSA) Millennium and the 3M FR-M40. The protective hood was the Joint Service Lightweight Integrated Suit Technology overcoat. The escape respirator was the Joint Service Chemical Environment Survivability Mask. The ballistic protective helmet was the MSA Advanced Combat Helmet – Commercial. The list was played three times for each of the test configurations listed in Table I.

Table 1. Test Configurations

Talker Headform	Listener Headform	PPE Impact
bareheaded	bareheaded	control
Millennium	bareheaded	speech transmission
FRM-40	bareheaded	speech transmission
bareheaded	protective hood	speech intelligibility
bareheaded	helmet	speech intelligibility
bareheaded	escape respirator	speech intelligibility

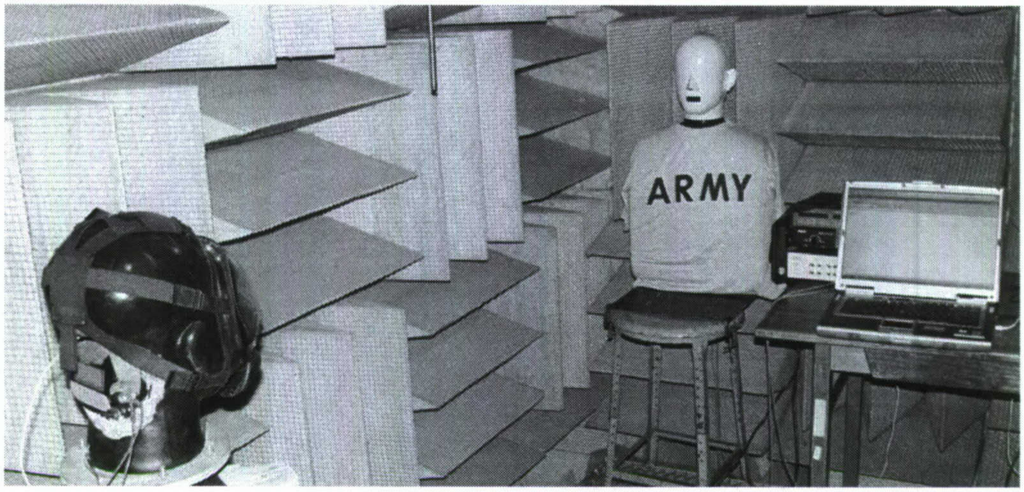

Figure 3. System Test Set Up Showing Talker and Listener Headforms in Anechoic Chamber

Three scores were determined for each trial: number of correctly identified high predictability words, number of correctly identified low predictability words, and total words identified correctly. Data for the speech transmission and speech intelligibility configurations were analyzed separately. An ANOVA was used to determine if there were statistically significant differences in scores among conditions while Tukey pairwise multiple comparisons were used to identify homogeneous subsets when there were differences.

3. RESULTS

3.1 Speech Recognition Software

Scores for each trial, performance ratings, and corresponding average sound levels are shown in Tables 2 and 3.

Table 2. MRT and SPIN Scores (out of 10) for Two Human Speakers (either bareheaded or wearing APR)

Speaker	List	Trial	MRT	SPIN
1	1	no mask	8	9
1	2	no mask	7	10
2	1	no mask	7	9
2	2	no mask	3	8
		Average	**6.3**	**9**
		SD	**2.2**	**0.8**
1	1	mask	3	9
1	2	mask	4	6
2	1	mask	5	8
2	2	mask	5.5*	9
		Average	**4.3**	**8**
		SD	**1.0**	**1.4**

*Volunteer skipped a sentence; the score was 5/9 (55%)

Table 3. Average Human Speaker Sound Levels

	Sound Level (dBA) MRT	Sound Level (dBA) SPIN
No mask	79.6 ± 6.6	76.6 ± 4.1
APR	78.7 ± 6.5	79.2 ± 4.3

3.2 System Evaluation

Scores for the SPIN test for each of the PPE configurations tested are shown in Tables 4 (speech transmission impacts) and Table 5 (speech intelligibility impacts). The "High" and "Low" scores are the number of correctly identified high and low predictability words out of a possible 25. The "Total" scores are the total number of words identified correctly out of 50. Superscripts indicate homogeneous groups.

Table 4. Speech Transmission Configurations

Talker Headform	Listener Headform	High	Low	Total
bareheaded	bareheaded	23 ± 1^A	21 ± 1^A	44 ± 1^A
Millennium	bareheaded	22 ± 1^A	18 ± 2^B	40 ± 2^B
FRM-40	bareheaded	18 ± 1^B	17 ± 1^B	35 ± 1^C

Note: Values are means ± standard deviations. Means with different letters are statistically significantly different at p = 0.05.

Table 5. Speech Intelligibility Configurations

Talker Headform	Listener Headform	High	Low	Total
bareheaded	bareheaded	23 ± 1^A	21 ± 1^A	44 ± 1^A
bareheaded	protective hood	19 ± 1^B	16 ± 3^B	35 ± 3^B
bareheaded	escape respirator	16 ± 2^C	14 ± 2^B	29 ± 1^C
bareheaded	helmet	24 ± 0^A	23 ± 1^A	47 ± 1^A

Note: Values are means ± standard deviations. Means with different letters are statistically significantly different at p = 0.05.

4. DISCUSSION

The MRT and SPIN scores shown in Table 2 demonstrate that the SPIN sentences provide high scores for the bareheaded condition and show reasonable degradations for mask wear. As the SPIN is easier to score and more closely replicates the process used by the software to identify words, the SPIN was selected as the speech intelligibility test for all subsequent work.

The SPIN scores shown in Tables 4 and 5 reflect changes in speech transmission and intelligibility caused by wearing PPE.

For the speech transmission configurations, the high, low, and total scores all provide useful information. The high score shows that the Millennium mask doesn't degrade speech for the tested conditions when context clues are used, but that the FRM-40 does. The low score demonstrates that both APRs degrade speech significantly when the speech is not predictable. Finally, the total score can be used to rank the APRs to indicate which one would be best for overall communications.

The high, low, and total scores are also useful for the speech intelligibility configurations. For all three scores, the control and helmet conditions were statistically the same. This was expected because the helmet does not cover the ears, but was useful for validating the method. For the low score, both the protective hood and escape hood degraded speech to the same degree. However, for the high and total scores, the protective hood impacted speech intelligibility less than the escape respirator hood. These scores would also be useful for ranking the impact of PPE on speech intelligibility.

5. CONCLUSIONS

An automated objective test system was developed to assess the impact of head-borne PPE on speech intelligibility and transmission. Preliminary results show that the system is capable of differentiating between different types of PPE. Improvements to the measurement technique are necessary to provide information useful for standards development. The speaker in the talker headform will be upgraded and more PPE tested. Further research will be necessary to investigate the impact of different speech sound levels, noise levels, and speaker-listener distances. These impacts must be correlated to human subject testing and must reflect operational performance requirements.

Blank

LITERATURE CITED

1. House, A.S.; Williams, C.E.; Hecker, M.H.L.; Kryter, K.D. Articulation Testing Methods: Consonantal Differentiation with a Closed-Response Set. *J. Acoust. Soc. Am.* **1965**, *37*, pp 158-166.

2. Kalikow, D.N.; Steven, K.N.; Elliott, L.L. Development of a Test of Speech Intelligibility in Noise Using Sentence Materials with Controlled Word Predictability. *J. Acoust. Soc. Am.* **1977**, *61*, pp 1337-1351.

3. *Determination of Communication Performance Test for Speech Conveyance and Intelligibility of Chemical Biological Radiological and Nuclear (CBRN) Full Facepiece Air-Purifying Respirator Standard Test Procedure (STP); Procedure No. CET-APRS-STP-CBRN-0313*; National Institute for Occupational Safety and Health (NIOSH), December 2005.

Blank

APPENDIX - WORD LISTS AND SUBJECT RESPONSES

MRT Set 1 (spoken word is in bold type)

kick, **lick**, sick, tick, wick, pick
neat, **beat**, seat, meat feat, heat
pun, **puff**, pup, pub, pus, puck
hook, shook, book, took, **cook**, look
lip, hip, dip, sip, rip, **tip**
rake, rate, ray, raze, race, **rave**
fang, bang, **hang**, sang, gang, rang
will, hill, kill, bill, fill, **till**
map, mat, **math**, mad, mass, man
pale, **sale**, bale, gale, male, tale

MRT Set 2 (spoken word is in bold type)

sack, sad, sap, sag, **sat**, sass
sit, sip, sill, sick, **sin**, sing
fold, sold, **gold**, hold, cold, told
but, bug, bus, **buff**, bun, buck
late, lake, **lay**, lame, lane, lace
run, bun, fun, sun, **nun**, gun
dust, gust, **must**, bust, just, rust
path, pack, pass, pat, **pad**, pan
dip, dim, **din**, dill, did, dig
fit, hit, bit, **sit**, kit, wit

SPIN sentences Set 1 (key word is in bold type)

1. The watchdog gave a warning **growl**.
2. She made the bed with clean **sheets**.
3. The old train was powered by **steam**.
4. He caught the fish in his **net**.
5. Close the window to stop the **draft**.
6. My T.V. has a twelve-inch **screen**.
7. The sandal has a broken **strap**.
8. The boat sailed along the **coast**.
9. Crocodiles live in muddy **swamps**.
10. The farmer harvested his **crop**.

SPIN sentences Set 2 (key word is in bold type)

1. All the flowers were in **bloom**.
2. She wore a feather in her **cap**.
3. The Admiral commands the **fleet**.
4. The beer drinkers raised their **mugs**.
5. He was hit by a poisoned **dart**.
6. The bread was made from whole **wheat**.
7. I made the phone call from a **booth**.
8. The cut on his knee formed a **scab**.
9. His boss made him work like a **slave**.
10. The farmer baled the **hay**.